



Draft Genome Sequence of *Desulfovibrio* sp. Strain CSMB_222, Isolated from Coal Seam Formation Water

 Andrew G. McLeish,^{a,b}  Paul Greenfield,^{b,c} David J. Midgley,^b  Ian T. Paulsen^a

^aDepartment of Molecular Sciences, Macquarie University, North Ryde, Australia

^bEnergy, Commonwealth Scientific and Industrial Research Organisation, Lindfield, Australia

^cDepartment of Biological Sciences, Macquarie University, North Ryde, Australia

ABSTRACT Subsurface coal seams contain microbial consortia with various taxa, each with a different role in the degradation of coal organic matter. This study presents the sequenced and annotated genome of *Desulfovibrio* sp. strain CSMB_222, a bacterium isolated from eastern Australian coal seams.

Coal seams harbor microbial consortia that catabolize organic matter in coal to methane (1, 2). This generation of biogenic methane has important environmental and economic significance. Methane has value as a transition fuel and a source of blue hydrogen (3). Thus, investigation of coal microbes and their roles in coal-to-methane transformations is critical to enhance methane yields and to improve our understanding of subsurface microbiology.

Formation water from Sydney Basin Well 2 was collected from the Bulli Seam (34°03'S, 150°46'E) at a depth of 650 m (4). This water was anoxically enriched at 40°C, in the dark, using Postgate C medium (5). Dilution to extinction was carried out until an axenic culture of a *Desulfovibrio* species was generated; 16S rRNA sequencing using Earth Microbiome Project primers (6) was used to assess culture purity. Genomic DNA was extracted using a PowerSoil DNA isolation kit (MO BIO Laboratories, Inc., USA), and library preparation was undertaken using the Nextera XT DNA library preparation kit (Illumina) according to the manufacturer's instructions. The resultant DNA library was sequenced on a HiSeq 2500 system, producing 150-bp paired-end reads (Macrogen, South Korea). The genome sequence data produced 5,291,357 raw reads, which were error corrected using Blue v2.1.4 (7) and assembled using SPAdes v3.13.2 (8), and full-length 16S rRNA was extracted using Kelpie v2.0.1 (9) (Genomic Information *Desulfovibrio* sp. CSMB_222; <https://www.doi.org/10.25919/m6yy-fc50>). The 16S rRNA sequence matched that of *Desulfovibrio* sp. strain CSMB_222 from the Coal Seam Microbiome (CSMB) data set (4). Phylogenetic analyses of 16S rRNA sequences revealed that *Desulfovibrio* sp. strain CSMB_222 was closely related (16S rRNA identity of >99%) to *Desulfovibrio vulgaris* Miyazaki F (GenBank accession number [CP001197.1](https://www.ncbi.nlm.nih.gov/nuccore/CP001197.1)), *Desulfovibrio termitidis* HI1 (GenBank accession number [NR_026255.1](https://www.ncbi.nlm.nih.gov/nuccore/NR_026255.1)), and *Desulfovibrio oxamicus* DSM1925 (GenBank accession number [NR_043567.1](https://www.ncbi.nlm.nih.gov/nuccore/NR_043567.1)), and distinct (16S rRNA identity of <96%) from *Desulfovibrio vulgaris sensu stricto*, a clade that includes the type strain *Desulfovibrio vulgaris* Hildenborough (GenBank accession number [NR_074446.1](https://www.ncbi.nlm.nih.gov/nuccore/NR_074446.1)) (Fig. 1).

The draft genome of *Desulfovibrio* sp. strain CSMB_222 was ~4.2 Mbp, comprising 219 contigs with a mean sequencing depth of ~50× and a GC content of 66.3%, and was 100% complete, with 0.58% contamination, according to CheckM v1.1.3 (10). The mean, median, and N_{50} contig lengths for the assembly were 19,017 bp, 11,315 bp, and 38,319 bp, respectively. The genome was annotated by Prokka v1.14.5 (11), which revealed 3,335 putative protein-coding genes, 55 tRNAs, and 2 rRNAs (<https://www.doi.org/10.25919/m6yy-fc50>). The genome was examined using TransportDB v2.0 (12) (<http://membranetransport.org/transportDB2>) and dbCAN v9.0 (13) (<http://bcb.unl.edu/dbCAN2>) to determine its transporters

Editor Catherine Putonti, Loyola University Chicago

© Crown copyright 2021. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Andrew G. McLeish, andrew.mcleish@csiro.au.

Received 17 July 2021

Accepted 6 November 2021

Published 2 December 2021

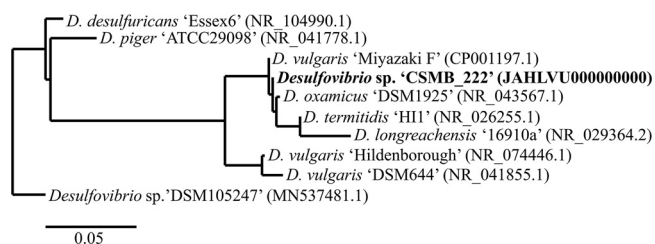


FIG 1 Phylogenetic tree based on 16S rRNA sequences, using a similarity matrix generated by PUP v1.0 analysis (21) (<https://github.com/PaulGreenfieldOz/WorkingDogs>). The similarity matrix was converted to the Newick format using hierarchical clustering (22) for phylogenetic visualization with TreeDyn v198.3 (23, 24). The numbers in parentheses are GenBank accession numbers.

and carbohydrate-active enzymes, respectively. The taxon possessed limited carbohydrate-active enzymes; in brief, 4 auxiliary activity (AA), 12 glycoside hydrolase (GH), and 4 carbohydrate esterase (CE) enzymes were detected, with no polysaccharide lyase (PL) enzymes. Only 3 carbohydrate-active enzymes with signal peptides were detected (GH13_33, GH23, and GH102). In terms of transporters, the taxon focused on the uptake of amino acids, peptides, and carboxylates, with an extensive array of ABC transporters, tripartite ATP-independent periplasmic transporters (TRAP-Ts), and tripartite tricarboxylate transporters (TTTs). Consistent with the paucity of carbohydrate-active enzymes, the strain possesses a single phosphotransferase system (PTS) and two ABC systems predicted for sugar uptake. Unless otherwise noted, default parameters were used for all software tools.

The *Desulfovibrio* sp. strain CSMB_222 genome contained two CRISPR arrays with 96 spacer sequences, suggesting that viral predation is common, as in many other taxa in coal seam subsurface environments (14–16). The role of this taxon in sulfate-poor coal seams in Australia is unknown; however, it seems likely to be engaged in syntrophy with hydrogenotrophic methanogens, similar to *Desulfovibrio desulfuricans* (17) and *Desulfovibrio vulgaris* (18–20).

Data availability. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession number [JAHLVU000000000](https://genbank.ncbi.nlm.nih.gov/GenBank/FASTA/JAHLVU000000000), BioProject number [PRJNA734375](https://bioinformatics.ncbi.nlm.nih.gov/bioproject/19490895), and BioSample number [SAMN19490895](https://bioinformatics.ncbi.nlm.nih.gov/biosample/SAMN19490895). The version described in this paper is version [JAHLVU010000000](https://bioinformatics.ncbi.nlm.nih.gov/biosample/SAMN19490895). The raw sequencing reads are available from the Sequence Read Archive (SRA) under BioProject number [PRJNA734375](https://bioinformatics.ncbi.nlm.nih.gov/bioproject/19490895).

ACKNOWLEDGMENTS

We acknowledge Nai Tran-Dinh and Silas Vick for assistance with anaerobic culturing and advice on experimental strategies, respectively, Debra Lock for logistical assistance, and Kaydy Pinetown for geological insight.

A.G.M. is funded by a postgraduate scholarship from Macquarie University, with additional support from the CSIRO Energy strategic research initiative.

REFERENCES

- Strąpoć D, Mastalerz M, Dawson K, Macalady J, Callaghan AV, Wawrik B, Turich C, Ashby M. 2011. Biogeochemistry of microbial coal-bed methane. *Annu Rev Earth Planet Sci* 39:617–656. <https://doi.org/10.1146/annurev-earth-040610-133343>.
- Ritter D, Vinson D, Barnhart E, Akob DM, Fields MW, Cunningham AB, Orem W, McIntosh JC. 2015. Enhanced microbial coalbed methane generation: a review of research, commercial activity, and remaining challenges. *Int J Coal Geol* 146:28–41. <https://doi.org/10.1016/j.coal.2015.04.013>.
- Das D, Veziroğlu TN. 2001. Hydrogen production by biological processes: a survey of literature. *Int J Hydrogen Energy* 26:13–28. [https://doi.org/10.1016/S0360-3199\(00\)00058-6](https://doi.org/10.1016/S0360-3199(00)00058-6).
- Vick SHW, Greenfield P, Tran-Dinh N, Tetu SG, Midgley DJ, Paulsen IT. 2018. The Coal Seam Microbiome (CSMB) reference set, a lingua franca for the microbial coal-to-methane community. *Int J Coal Geol* 186:41–50. <https://doi.org/10.1016/j.coal.2017.12.003>.
- Postgate JR. 1984. The sulphate reducing bacteria, p 32. Cambridge University Press, Cambridge, England.
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6:1621–1624. <https://doi.org/10.1038/ismej.2012.8>.
- Greenfield P, Duesing K, Papanicolaou A, Bauer DC. 2014. Blue: correcting sequencing errors using consensus and context. *Bioinformatics* 30:2723–2732. <https://doi.org/10.1093/bioinformatics/btu368>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Greenfield P, Tran-Dinh N, Midgley D. 2019. Kelpie: generating full-length “amplicons” from whole-metagenome datasets. *PeerJ* 6:e6174. <https://doi.org/10.7717/peerj.6174>.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells,

- and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
11. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
 12. Elbourne LDH, Tetu SG, Hassan KA, Paulsen IT. 2017. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res* 45:D320–D324. <https://doi.org/10.1093/nar/gkw1068>.
 13. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 40:W445–W451. <https://doi.org/10.1093/nar/gks479>.
 14. Daly RA, Borton MA, Wilkins MJ, Hoyt DW, Kountz DJ, Wolfe RA, Welch SA, Marcus DN, Trexler RV, MacRae JD, Krzycki JA, Cole DR, Mouser PJ, Wrighton KC. 2016. Microbial metabolisms in a 2.5-km-deep ecosystem created by hydraulic fracturing in shales. *Nat Microbiol* 1:16146. <https://doi.org/10.1038/nmicrobiol.2016.146>.
 15. Vick SHW, Greenfield P, Tetu SG, Midgley DJ, Paulsen IT. 2019. Genomic and phenotypic insights point to diverse ecological strategies by facultative anaerobes obtained from subsurface coal seams. *Sci Rep* 9:16186. <https://doi.org/10.1038/s41598-019-52846-7>.
 16. Vick SHW, Greenfield P, Willows RD, Tetu SG, Midgley DJ, Paulsen IT. 2020. Subsurface *Stappia*: success through defence, specialisation and putative pressure-dependent carbon fixation. *Microb Ecol* 80:34–46. <https://doi.org/10.1007/s00248-019-01471-y>.
 17. Cord-Ruwisch R, Lovley DR, Schink B. 1998. Growth of *Geobacter sulfurreducens* with acetate in syntrophic cooperation with hydrogen-oxidizing anaerobic partners. *Appl Environ Microbiol* 64:2232–2236. <https://doi.org/10.1128/AEM.64.6.2232-2236.1998>.
 18. Bryant MP, Campbell LL, Reddy CA, Crabill MR. 1977. Growth of *Desulfovibrio* in lactate or ethanol media low in sulfate in association with H₂-utilizing methanogenic bacteria. *Appl Environ Microbiol* 33:1162–1169. <https://doi.org/10.1128/aem.33.5.1162-1169.1977>.
 19. Scholten JC, Culley DE, Brockman FJ, Wu G, Zhang W. 2007. Evolution of the syntrophic interaction between *Desulfovibrio vulgaris* and *Methanosarcina barkeri*: involvement of an ancient horizontal gene transfer. *Biochem Biophys Res Commun* 352:48–54. <https://doi.org/10.1016/j.bbrc.2006.10.164>.
 20. Walker CB, He Z, Yang ZK, Ringbauer JA, He Q, Zhou J, Voordouw G, Wall JD, Arkin AP, Hazen TC, Stolyar S, Stahl DA. 2009. The electron transfer system of syntrophically grown *Desulfovibrio vulgaris*. *J Bacteriol* 191:5793–5801. <https://doi.org/10.1128/JB.00356-09>.
 21. Sutcliffe B, Chariton AA, Harford AJ, Hose GC, Stephenson S, Greenfield P, Midgley DJ, Paulsen IT. 2018. Insights from the genomes of microbes thriving in uranium-enriched sediments. *Microb Ecol* 75:970–984. <https://doi.org/10.1007/s00248-017-1102-z>.
 22. Boc A, Diallo AB, Makarenkov V. 2012. T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucleic Acids Res* 40:W573–W579. <https://doi.org/10.1093/nar/gks485>.
 23. Chevenet F, Brun C, Bañuls A-L, Jacq B, Christen R. 2006. TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* 7:439. <https://doi.org/10.1186/1471-2105-7-439>.
 24. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard J-F, Guindon S, Lefort V, Lescot M, Claverie J-M, Gascuel O. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 36:W465–W469. <https://doi.org/10.1093/nar/gkn180>.